Cyber-attacks detection in industrial systems using artificial intelligence-driven methods

# Wu Wang; Fouzi Harrou; Benamar Bouyeddou et al.

## 2022

Results suggested that using ensemble learning models and exploiting their capacity to provide precise detection significantly reducing the number of missed detections and false alarms to reach high cyber-attack detection performance

# Abstract

Modern industrial systems and critical infrastructures are constantly exposed to malicious cyber-attacks that are challenging and difficult to identify. Cyber-attacks can cause severe economic losses and damage the attacked system if not detected accurately and timely. Therefore, designing an accurate and sensitive intrusion detection system is undoubtedly necessary to ensure the productivity and safety of industrial systems against cyber-attacks. This paper first introduces a stacked deep learning method to detect malicious attacks in SCADA systems. We also consider eleven machine learning models, including the Xtreme Gradient Boosting (XGBoost), Random forest, Bagging, support vector machines with different kernels, classification tree pruned by the minimum cross-validation and by 1-standard error rule, linear discriminate analysis, conditional inference tree, and the C5.0 tree. Real data sets with different kinds of cyber-attacks from two laboratory-scale SCADA systems, gas pipeline and water storage tank systems, are employed to evaluate the performance of the investigated methods. Seven evaluation metrics have been used to compare the investigated models (accuracy, sensitivity, specificity, precision, recall, F1-score, and area under curve, or AUC). Overall, results show that the XGBoost approach achieved superior detection performance than all other investigated methods. This could be due to its desirable characteristics to avoid overfitting, decreases the complexity of individual trees, robustness to outliers, and invariance to scaling and monotonic transformations of the features. Unexpectedly, the deep learning models are not providing the best performance in this case study, even with their extended capacity to capture complex features interactions.

# Study subjects

**225000 customers**

The damage was minimized due to the hackers’ limited knowledge of ICS/SCADA systems or the intention to do any harm, but it could have severe consequences. In December 2015, a cyber-attack on the Ukraine power grid caused circuit breakers at 30 substations to trip, cutting power to around 225,000 customers [13]. A denial of service (DoS) targeted the telephone system and communication network, making the call center unavailable to customers**In December 2015, a cyber-attack on the Ukraine power grid caused circuit breakers at 30 substations to trip, cutting power to around 225,000 customers [13]**

# Findings

Results revealed that the XGBoost model yields higher detection performance than other investigated models and outperforms the state-of-the-art methods

From Table 3, we can see the p-values for comparing Net1, Net4, and Net5 with the stacked deep learning model are small, and the test is rejected under a significance level of 0.1, meaning that the stacked deep learning model performs significantly better than individual deep learning models in terms of accuracy

If the observation is declared an attack, it is classified into one of seven attack classes in the second stage. After that, another sub-attack classification is accomplished to find the sub-attack type. This three-stage-based intrusion detection scheme achieved an accuracy of 92.06%

The one-class detection methods, including OCC-eSNN, OCC-SVM, OCC-CDCPE [30] exhibit the closest performance to ensemble models (i.e., XGBoost and RF) by obtaining a detection accuracy of 98.08%, 98.01%, 96.75%, respectively

For the study reported in [51], detection results using the combined K-means-CAE approach are relatively below than one-class-based and ensemble models by reaching an accuracy of 96.59%

The HMM-based detector [29] have relatively lower performance compared to the other models (an accuracy of 93.36%) mainly because it is not efficient in dealing with such complicated multivariate data

Although XGboost and RF models have superior detection performance than the stacked deep learning model in this work; it is hard to say that ensemble learning models dominate the deep learning models. This is because we expect that deep learning can achieve improved performance once larger datasets are used

# Scholarcy Highlights

* The latest (Fourth) generation of SCADA (Supervisory Control and Data Acquisition) systems integrate different internet and networks technologies, such as Cloud computing, Internet of things (IOT), big data, and web services
* The success behind machine learning methods such as random forest and Xtreme Gradient Boosting (XGBoost) is largely due to their robustness to outliers
* The performances of eleven machine learning models, including the Xtreme Gradient Boosting (XGBoost), Random forest, Bagging, support vector machines with different kernels, classification tree pruned by the minimum crossvalidation and by 1-standard error rule, linear discriminate analysis, conditional inference tree, and the C5.0 tree, are investigated in detecting attacks in SCADA systems
* Results suggested that using ensemble learning models (XGBoost and random forest (RF)) and exploiting their capacity to provide precise detection significantly reducing the number of missed detections and false alarms to reach high cyber-attack detection performance
* XGboost and RF models have superior detection performance than the stacked deep learning model in this work; it is hard to say that ensemble learning models dominate the deep learning models
* This is because we expect that deep learning can achieve improved performance once larger datasets are used
* How to design and implement a stacked deep learning model for cyber-attack detection that is robust to outliers is a challenge that we leave to future research

# Scholarcy Summary

## Introduction

The latest (Fourth) generation of SCADA (Supervisory Control and Data Acquisition) systems integrate different internet and networks technologies, such as Cloud computing, Internet of things (IOT), big data, and web services.

Such systems can provide real-time notifications, allowing administrators to use a variety of platforms to access the system and execute many new complex control algorithms [1].

Despite these advantages, this growing connectivity has made them more vulnerable to different forms of cyber-attacks.

Several security mechanisms, including intrusion detection systems and cryptography, are adopted to secure and protect industrial systems from security threats [7]

## Methods

The deep learning and stacked deep learning methods, and support vector machine based methods are inferior to the tree based ensemble methods in discriminating cyber-attacks in this study.

The accuracy and F1-score are 0.968 and 0.956, respectively, for the stacked deep learning method, 0.972 and 0.962, respectively, for the support vector machine using the polynomial kernel, which are smaller than those of the XGBoost method.

It can be observed that results testify the superior discrimination capacity of the ensemble learning models (i.e., XGBoost, RF, and Bagging) compared to the other models, including deep learning models.

Based on a significance level of 0.1, the accuracy of the stacked deep learning model is significantly better than that of Net2, Net4 and Net5

## Results

This section is dedicated to verify the performance of the considered models using datasets from two testbeds: the gas pipeline and the storage tank testbeds from the Mississippi State University SCADA (11) Laboratory [47].

The deployed physical process and control systems include gas pipeline and water tank storage in which normal operations are altered by different type of attacks (Fig. 3).

The effectiveness of the considered cyberattack detectors to uncover cyber-attacks is investigated using SCADA datasets collected from the testbeds of SCADA Laboratory of Mississippi State University.

The gas pipeline testbed mimics the behavior of a gas pipeline in transporting petroleum products to the market, while the storage tank testbed reproduces the oil storage tanks found in the petrochemical industry

This testbed includes principally sensors and actuators, a communication network, and supervisory control systems.

About 28 scenarios were created, including reconnaissance attacks, response injection attacks, command injection attacks, and denial of Service attacks [47,48]

## Conclusion

With the extended use of information and communication technologies in SCADA systems, cyberattacks against industrial systems have significantly increased.

The performances of eleven machine learning models, including the Xtreme Gradient Boosting (XGBoost), Random forest, Bagging, support vector machines with different kernels, classification tree pruned by the minimum crossvalidation and by 1-standard error rule, linear discriminate analysis, conditional inference tree, and the C5.0 tree, are investigated in detecting attacks in SCADA systems.

To this end, measurements from a gas pipeline and water storage tank testbeds are used to verify the performance of the considered machine learning and deep learning models.

Results suggested that using ensemble learning models (XGBoost and RF) and exploiting their capacity to provide precise detection significantly reducing the number of missed detections and false alarms to reach high cyber-attack detection performance

# Builds on previous work

Deep learning methods is successful in various fields such as speech recognition [40], natural language processing [41], and computer vision [42]. **We implement networks with different number of hidden neurons to cover small to large neural networks and compare their prediction accuracy in the experiment**.

Stack generalization [43] is a popular strategy widely adopted in the machine learning community to boost the performance of machine learning algorithms. While a single deep neural network may overfit the training data and **does not have enough power for discriminating cyber-attacks, we adopt an ensemble learning strategy and develop a stacked deep learning method**.

From Table 5, we noticed that the state-of-the-art methods are providing low classification performance for MSCI, CMRI, and MPCI attack types compared to the other attacks. **This is mainly because of the noisiness of data related to these attacks since they are related to the physical processes that show generally noisy behavior** [52]

# Contributions

With the extended use of information and communication technologies in SCADA systems, cyberattacks against industrial systems have significantly increased. The task of conventional IDSs becomes very challenging due to the several vulnerabilities in SCADA systems and the variety of cyberattacks. Miss-detection of cyberattacks in critical systems can lead to severe economic and safety consequences. This work explores the detection potential of machine learning and deep learning to support IDSs in detecting cyberattacks. At first, this study investigated the feasibility of a stacked deep learning scheme to uncover malicious attacks in SCADA systems. Moreover, the performances of eleven machine learning models, including the Xtreme Gradient Boosting (XGBoost), Random forest, Bagging, support vector machines with different kernels, classification tree pruned by the minimum crossvalidation and by 1-standard error rule, linear discriminate analysis, conditional inference tree, and the C5.0 tree, are investigated in detecting attacks in SCADA systems. To this end, measurements from a gas pipeline and water storage tank testbeds are used to verify the performance of the considered machine learning and deep learning models. This comparison study generates that the XGBoost-anomaly detection approach achieved superior detection performance than all other investigated models. The RF scheme also achieves comparable performances. Results suggested that using ensemble learning models (XGBoost and RF) and exploiting their capacity to provide precise detection significantly reducing the number of missed detections and false alarms to reach high cyber-attack detection performance.

# Future work

Section 5 concludes this study and sheds light on potential future research lines.

This is because we expect that deep learning can achieve improved performance once larger datasets are used. How to design and implement a stacked deep learning model for cyber-attack detection that is robust to outliers is a challenge that we leave to future research.

These machine learning models are constructed using labeled training data, while obtaining this prior knowledge on all the existing types of attacks is not an easy task, with the generation of advanced and sophisticated attacks every day. An interesting direction for future work is the design of unsupervised anomaly detection approaches by amalgamating unsupervised deep learning models as features extractors, such as deep generative models, with the sensitivity of statistical monitoring charts, such as Generalized Likelihood Ratio Test [58].